

Lévy-Khintchine random matrices and the Poisson weighted infinite skeleton tree

Paul Jung

Department of Mathematics, University of Alabama Birmingham

August 5, 2014

Abstract

We study a class of Hermitian random matrices which includes Wigner matrices, heavy-tailed random matrices, and sparse random matrices such as adjacency matrices of Erdős-Rényi random graphs with $p_n \sim \frac{1}{n}$. Our $n \times n$ random matrices have real entries which are i.i.d. up to symmetry. The distribution of entries depends on n , and we require sums of rows to converge in distribution; it is then well-known that the limit distribution must be infinitely divisible.

We show that a limiting empirical spectral distribution (LSD) exists, and via local weak convergence of associated graphs, the LSD corresponds to the spectral measure associated to the root of a graph which is formed by connecting infinitely many Poisson weighted infinite trees using a backbone structure of special edges called “cords to infinity”. One example covered by the results are matrices with i.i.d. entries having infinite second moments, but normalized to be in the Gaussian domain of attraction. In this case, the limiting graph is \mathbb{N} rooted at 1, so the LSD is the semi-circle law. The results also extend to self-adjoint complex matrices and Wishart matrices.

Keywords: Empirical spectral distribution, Wigner matrices, Lévy matrices, heavy-tailed random matrices, sparse random matrices, Erdős-Rényi graph, local weak convergence, cavity method.

Contents

1	Introduction	2
1.1	Main results	4
2	A replacement procedure for cords to infinity	10
3	The moment method	11
4	From local weak convergence to spectral convergence	16
5	Proofs of the main results	21
A	Some additional tools	25

1 Introduction

This paper jointly studies the limiting spectral distributions (LSD) for three classes of Hermitian random matrices that have appeared in the literature. The first class of random matrices are classic Wigner matrices introduced in the seminal work of their namesake, [Wig55]. The literature on this class of random matrices is overwhelmingly abundant (see [AGZ10, BS10, Tao12]).

The second class of matrices are adjacency matrices of Erdős-Rényi random graphs on n vertices whose edges are present with probability proportional to $1/n$. The analysis of the LSD in the context of random matrices seems to have started in [RB88]. These matrices are called *sparse¹ random matrices*, and they can be considered a Poissonian variation of Wigner matrices. The LSD of sparse random matrices was analyzed using the “moment method” in [Rya98, BG01, KSV04, Zak06], and using the “resolvent method” in [KSV04]. An insightful modification of the latter approach led to improved results in [BL10]. (see also [Küh08] for references in the physics literature).

Finally, the third class of random matrices are formed from properly normalized heavy-tailed entries and, following [BAG08], we call them *heavy-tailed random matrices*. These are also known in the physics literature as Lévy matrices or Wigner-Lévy matrices, and they were introduced by Cizeau and Bouchaud in [CB94]. Later, they were studied more rigorously in [Sos04, BAG08, BCC11a]. These matrices are *not* to be confused with *free* Lévy matrices [BG05, BJN⁺07].

In each of the three classes of matrices above, the entries are i.i.d. up to self-adjointness, although the distributions may differ for different n . In order to obtain non-trivial LSDs, a proper rescaling or change in distribution is needed as $n \rightarrow \infty$ (such rescaling is often implicit in the formulation). After respectively rescaling, if one sums all the entries in a single row or column and takes $n \rightarrow \infty$, then one obtains a Gaussian, Poisson, or stable distribution in each of the respective classes. These are all examples of infinitely divisible distributions which suggests that all three classes of matrices can many times be thought of under this umbrella, and various papers (for example [Rya98, Sec. 3.1]) have done exactly that. More recently, [BGGM13] establishes a functional central limit theorem for the Cauchy-Stieltjes transforms of the LSDs of all three classes, and [Mal12] studies the joint LSDs of a pair of independent ensembles in these three classes using algebraic techniques inspired by free probability.

Here, we also view these three classes as examples from this larger class of matrix ensembles characterized by the Lévy-Khintchine formula, and in particular, the matrices are viewed as

¹Here, the random number of non-zero entries remains bounded in distribution as $n \rightarrow \infty$. The term “sparse” sometimes refers to dilute random matrices for which the order of non-zero entries is $o(n)$.

(weighted) adjacency matrices. As was done in the heavy-tailed setting in [BCC11a], our main objective is to equate the LSD of the limiting adjacency operator with the spectral measure associated to the root (or vacuum state) vector in $L^2(V)$ where V is the vertex set of the limiting graph in the sense of local weak convergence (see below). This allows for further analysis of the LSD using the recursive structure of the limiting graph.

The ensembles we consider have i.i.d. complex entries for each n , up to self-adjointness, with zeros on the diagonal. It is well-known that any weak limit of sums of rows must be infinitely divisible in \mathbb{R}^2 . Actually, the “identically distributed” condition may be weakened to require only that the moduli of the entries are identically distributed. In this weakened form one still has that the sums of the square-moduli of entries, i.e. the Euclidean norm of a row as a vector in \mathbb{R}^{2n} , converge in distribution to a positive law which is the marginal distribution of a Lévy subordinator.

In particular, recall (see [Kyp06] or [Kal02]) that a probability distribution μ on \mathbb{R} is infinitely divisible with distribution $ID(\sigma^2, b, \Pi)$ and Lévy exponent Ψ ,

$$e^{\Psi(\theta)} := \int_{\mathbb{R}} e^{i\theta x} \mu(dx) \quad \text{for } \theta \in \mathbb{R},$$

if and only if there exists a triplet of characteristics (σ^2, b, Π) such that

$$\Psi(\theta) := -\frac{1}{2}\theta^2\sigma^2 + i\theta b + \int_{\mathbb{R}} e^{i\theta x} - 1 - \frac{i\theta x}{1+x^2} \Pi(dx), \quad (1)$$

where $\sigma^2 \geq 0$, $b \in \mathbb{R}$, and $\Pi(dx)$ concentrates on $\mathbb{R} \setminus \{0\}$ and satisfies

$$\int_{\mathbb{R}} (1 \wedge |x|^2) \Pi(dx) < \infty. \quad (2)$$

If μ concentrates on $(0, \infty)$ then the exponent corresponds to the subordinator characteristics (σ^2, Π_s) (which is equivalent to the triplet $(0, \sigma^2, \Pi_s)$) and takes the simplified form

$$\Psi_s(\theta) := i\theta\sigma^2 + \int_{(0, \infty)} (e^{i\theta x} - 1) \Pi_s(dx), \quad (3)$$

where $\Pi_s(dx)$ also concentrates on $(0, \infty)$, but instead of (2), it satisfies

$$\int_{(0, \infty)} (1 \wedge x) \Pi_s(dx) < \infty.$$

Here, the s subscript indicates the *subordinator* form of the Lévy exponent.

We say a sequence of $n \times n$ random matrices $(\mathcal{C}_n)_{n \in \mathbb{N}}$ is a **Lévy-Khintchine random matrix ensemble** with characteristics $(\sigma^2, 0, \Pi)$ if for each n , the moduli of entries $\mathcal{C}_n(j, k) = \bar{\mathcal{C}}_n(k, j)$, $j \neq k$ are i.i.d. (up to self-adjointness, with zeros on the diagonal) and the

$$\text{weak limit } \lim_{n \rightarrow \infty} \sum_{k=1}^n \pm |\mathcal{C}_n(1, k)| \text{ is infinitely divisible with characteristics } (\sigma^2, 0, \Pi), \quad (4)$$

where the signs \pm are independent Rademacher random variables (independent also from \mathcal{C}_n). This implies that Π is a symmetric measure. Note also that any deterministic “drift” in the original entries must be of order $1/n$ due to the i.i.d. condition together with weak convergence of the Euclidean norm of a row. Thus, any such drift does not contribute to the limiting sum due to the random signs

(the limiting sum of the randomly signed drift converges in probability to zero since it has zero mean and variance of order $1/n$). An equivalent form of the above is that the

$$\text{weak limit } \lim_{n \rightarrow \infty} \sum_{k=1}^n |\mathcal{C}_n(1, k)|^2 \text{ is infinitely divisible with subordinator characteristics } (\sigma^2, \Pi_s). \quad (5)$$

We note that by standard arguments, one could set the diagonal elements to any real number which converges to 0 fast enough, and this would not affect the LSD. For the sake of simplicity, we will always set diagonal entries to zero.

In the context of Lévy processes, the three components of the triplet (σ^2, b, Π) correspond to a Brownian component, a drift component, and a jump component (with possibly additional “compensating drift”), respectively. We will see in our context that σ^2 corresponds to a Wigner component, the drift component is inconsequential since by using the random signs it becomes 0 (cf. [BAG08, Remark 1.9]), and the Lévy measure Π generalizes both heavy-tailed and sparse random matrices.

Let $(\mathcal{A}_n)_{n \in \mathbb{N}}$ denote an ensemble which satisfies the above conditions except it does not require the condition of self-adjointness, $\mathcal{C}_n(j, k) = \bar{\mathcal{C}}_n(k, j)$. We call this a non-Hermitian Lévy-Khintchine random matrix ensemble. Using a standard bipartization/Hermitization method², our results extend to the LSD of Wishart matrices $(\mathcal{A}^* \mathcal{A}_n)_{n \in \mathbb{N}}$ or equivalently to the limiting empirical singular value distribution for $(\mathcal{A}_n)_{n \in \mathbb{N}}$.

1.1 Main results

For a given Lévy-Khintchine ensemble, let $\{\lambda_j\}_{j=1}^n$ denote the eigenvalues of the n th matrix in the sequence. The empirical spectral distribution (ESD) is defined as

$$\mu_{\mathcal{C}_n} := \frac{1}{n} \sum_{j=1}^n \delta_{\lambda_j}. \quad (6)$$

Theorem 1.1 (Existence of the LSD). *For any Lévy-Khintchine random matrix ensemble $(\mathcal{C}_n)_{n \in \mathbb{N}}$ with characteristics $(\sigma^2, 0, \Pi)$ (or alternatively with subordinator characteristics (σ^2, Π_s)), there exists a symmetric nonrandom probability measure $\mu_{\mathcal{C}_\infty}$ to which the ESDs $\mu_{\mathcal{C}_n}$ weakly converge, almost surely, as $n \rightarrow \infty$. In other words,*

$$\mathbf{P} \left(\lim_{n \rightarrow \infty} \langle \mu_{\mathcal{C}_n}, f \rangle = \langle \mu_{\mathcal{C}_\infty}, f \rangle \text{ for all bounded continuous } f \right) = 1. \quad (7)$$

Moreover, the limiting measure $\mu_{\mathcal{C}_\infty}$ has bounded support if and only if Π is trivial.

An extension of the above result to the singular values of $(\mathcal{A}_n - zI_n)_{n \in \mathbb{N}}$ in the spirit of [DS07] follows by way of Theorem 2.1 in [BCC11b] (see also [FZ97]). This gives us the following corollary.

Corollary 1.2 (LSD for Wishart ensembles). *Suppose $(\mathcal{A}_n)_{n \in \mathbb{N}}$ is a non-Hermitian Lévy-Khintchine ensemble with characteristics $(\sigma^2, 0, \Pi)$. The LSD, ν_∞ , of the Wishart ensemble $(\mathcal{A}^* \mathcal{A}_n)_{n \in \mathbb{N}}$ exists and is given by*

$$\nu_\infty(B) = \mu_{\mathcal{C}_\infty} \{x : x^2 \in B\}$$

where $\mu_{\mathcal{C}_\infty}$ is the LSD from Theorem 1.1 for the Hermitian ensemble with the same characteristics.

²This method has appeared in the physics literature (see [FZ97] and the references therein) and is discussed in the texts [AGZ10, Tao12].

In the case where Π has exponential moments, an extension of the standard moment method is enough to handle the proof of Theorem 1.1, and in Section 3 we do just that under the slightly stronger assumption that Π has bounded support. Under this latter assumption, one can also apply use Voiculescu's asymptotic freeness theorem to conclude existence (see Section 5). When Π has some moments which are infinite and $\sigma = 0$, the proof follows by generalizing insightful *local weak convergence* arguments of [BL10, BCC11a] (see Section 4). To extend this to the general case, we combine the local weak convergence arguments with a generalized moment method, which applies to both the sparse and Wigner case, with tail truncation arguments.

As a by-product of local weak convergence, one can view the LSD of the random matrix ensembles as the spectral measure of a weighted adjacency operator, at the root vector, of some new infinite graph. For ensembles with characteristics $(0, 0, \Pi)$, this idea is again a generalization of arguments in [BCC11a]. However, when $\sigma > 0$ a non-trivial generalization of Aldous' Poisson weighted infinite tree which we call a **Poisson weighted infinite skeleton tree (PWIST)** is required.

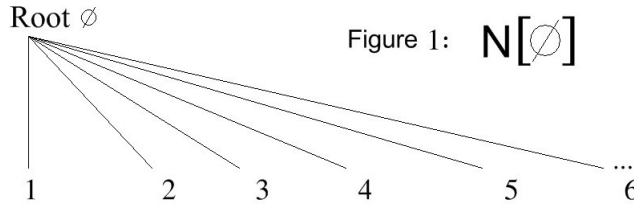
The idea of local weak convergence was introduced by Benjamini and Schramm [BS01] and further developed by Aldous and Steele [AS04]. Aldous and Steele describe the technique as finding “a new, infinite, probabilistic object whose local properties inform us about the limiting properties of a sequence of finite problems.” When the limiting object has a tree structure, local weak convergence provides a general framework to make the *cavity method* in physics rigorous. In our context, the cavity method was used in [CB94] and our new infinite object (with a tree structure) generalizes Aldous' Poisson infinite weighted trees (PWIT) by adding to it “cords” of infinite length which connect to independent copies of other PWITs. These cords form a *backbone structure* for a collective object which we refer to as a PWIST.

Let us first recall the definition of the PWIT(λ_Π). Start with a single root vertex \emptyset with an infinite number of (first generation) children indexed by \mathbb{N} . The weight on the edge to the k th child is the k th arrival (ordered by absolute value) of a Poisson process on $\mathbb{R} \setminus \{0\}$ with some intensity λ . In our situation the intensity λ_Π is derived from the measure Π on $\mathbb{R} \setminus \{0\}$ by inverting:

$$\lambda_\Pi\{x : 1/x \in B\} := \Pi(B) \quad (8)$$

For example, if $\Pi(dx)$ is absolutely continuous with density $f_\Pi(x)dx$ then $\lambda_\Pi(dx)$ is also absolutely continuous with density $x^{-2}f_\Pi(1/x)dx$ where x^{-2} is the change-of-measure factor.

If G has a root at \emptyset we write $G[\emptyset]$ for the rooted graph with (random) weights assigned to each edge. Slightly abusing notation, we denote the subgraph of a PWIT(λ_Π) formed by the root \emptyset , its children, and the weighted edges in between, by $\mathbb{N}[\emptyset]$.



We continue now with other generations. Every vertex v in generation $g \geq 1$ is given an infinite number of children indexed by \mathbb{N} forming the subgraph $\mathbb{N}[v]$. Thus the vertex set is

$$\mathbb{N}^F := \bigcup_{g \geq 0} \mathbb{N}^g \quad (9)$$

where $\mathbb{N}^0 = \emptyset$. The weights on edges to children in generation $g + 1$, from some fixed vertex v in generation g , are found by repeating the procedure for the weights in the first generation, namely according to the points of an independent Poisson random measure with intensity $\lambda_\Pi(dx)$.

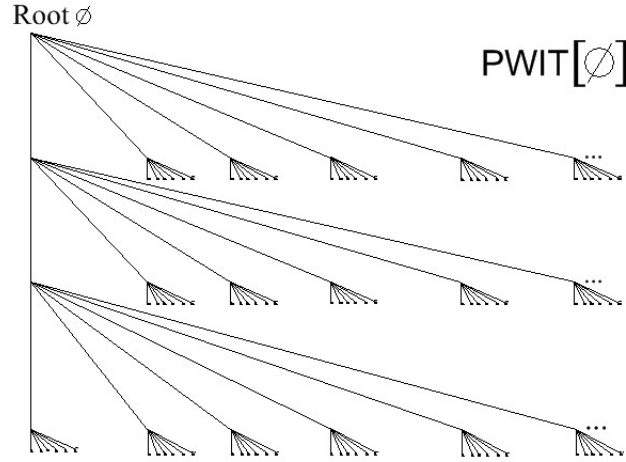


Figure 2: Each  represents a copy of the PWIT.

Weights on offspring edges from any vertex
are determined by a Poisson process.

The PWIST will depend on both characteristics σ^2 and Π (via λ_Π). To construct a $\text{PWIST}(\sigma, \lambda_\Pi)$, we start with a single $\text{PWIT}(\lambda_\Pi)[\emptyset]$ rooted at \emptyset and, for each vertex v of $\text{PWIT}(\lambda_\Pi)[\emptyset]$, we create a new vertex ∞_v which is the root of a new independent $\text{PWIT}(\lambda_\Pi)[\infty_v]$. We draw an edge from v to ∞_v for each v and assign this edge a nonrandom weight of

$$1/\sigma \in (0, \infty]. \quad (10)$$

Next, we create a new independent $\text{PWIT}(\lambda_\Pi)[\infty_u]$ for each vertex u of each $\text{PWIT}(\lambda_\Pi)[\infty_v]$, and draw an edge with weight $1/\sigma$ between u and ∞_u . We continue this procedure ad infinitum. If we also identify ∞_v with the integer 0 so that by concatenation, ∞_v is written $v0$, then we may write the vertex set of a $\text{PWIST}(\sigma, \lambda_\Pi)$ as

$$\mathbb{N}_0^F := \bigcup_{g \geq 0} \mathbb{N}_0^g \quad (11)$$

where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and by concatenation we write $v = v_1 v_2 \cdots v_g \in \mathbb{N}_0^g$. As can be seen in the figure below, edges with the weight $1/\sigma$ connect infinitely many PWITs with a backbone structure in order to form a PWIST.

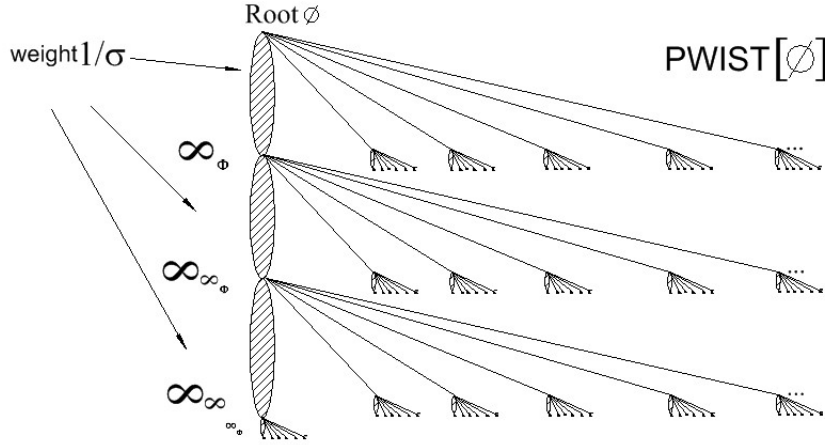


Figure 3: Each  represents a copy of the PWIST.

Weights on cords to infinities are deterministic. All other weights are random and determined by Poisson processes as before.

Our next theorem justifies the choice (10) for the weight on the edge between v and ∞_v . Let us however give a brief heuristic explanation as to why this is the correct weight to assign to this edge. First of all, identify each weight with its absolute value so that all weights are thought of as *nonnegative conductances*. Now, if $\sigma = 0$, then the connected graph containing the root \emptyset is simply a $\text{PWIT}(\lambda_\Pi)$ with the weights on edges representing nonnegative conductances. If $\sigma > 0$, we use the interpretation that v and ∞_v are infinitely far apart, but also that there are infinitely many parallel edges (or a multi-edge) between v and ∞_v . Since distance is equivalent to resistance on electrical networks and resistance is the reciprocal of conductance, the conductance of each parallel edge is zero; however, their collective *effective* conductance may be greater than 0. We can thus identify the multiple parallel edges with a *single* edge between v and ∞_v called a **cord to infinity** with effective resistance $1/\sigma$.

Let us now consider a random weighted adjacency matrix \mathcal{C}_{G_n} associated to a complete **rooted geometric graph** (see Section 4 for definitions) $G_n = G_n[\emptyset] = (V_n, E_n, \mathcal{R}_n)$ where $V_n = \{1, \dots, n\}$ and \mathcal{R}_n are the (possibly signed) random weights/lengths/resistances of the edges E_n . We refer to such a real-valued matrix as a *random conductance matrix* with entries given simply by the reciprocals of the signed resistances:

$$\mathcal{C}_{G_n}(j, k) := \frac{1}{\mathcal{R}_n(j, k)}. \quad (12)$$

When a sequence of random conductance matrices satisfies (4) or (5), it forms a Lévy-Khintchine random matrix ensemble.

This notion generalizes to a **random conductance operator** on $L^2(G_\infty) \equiv L^2(V_\infty)$ for an infinite weighted graph $G_\infty = (V_\infty, E_\infty, \mathcal{R}_\infty)$. Let the core $\mathcal{D}_{\text{fs}} \subset L^2(V_\infty)$ be the set of vectors with finite support, i.e., all finite linear combinations of the basis vectors e_v which are 1 at v and 0 elsewhere. We consider the operator on \mathcal{D}_{fs} which is defined by

$$\mathcal{C}_{G_\infty}(u, v) = \langle e_u, \mathcal{C}_{G_\infty} e_v \rangle := \begin{cases} 1/\mathcal{R}_\infty(u, v) & \text{if } u \sim v \\ 0 & \text{otherwise.} \end{cases} \quad (13)$$

This operator is closable as a graph in $L^2(V_\infty) \times L^2(V_\infty)$ since it is symmetric, i.e., Hermitian and

densely defined [WS80, Thm 5.4]. Abusing notation we also denote its unique closure by \mathcal{C}_{G_∞} . In particular, we will see that the closure is self-adjoint. In the case where G_∞ is a PWIST(σ, λ_Π), by (8), the conductances are given by the points of a Poisson random measure with symmetric intensity $\Pi(dx)$ on $\mathbb{R} \setminus \{0\}$.

Now, recall [RS80, Sec. VII.2 and VIII.3] that the spectral measure μ_φ of a self-adjoint operator \mathcal{C} associated to the vector φ is defined by the relation

$$\langle \varphi, f(\mathcal{C})\varphi \rangle =: \int_{\mathbb{R}} f(x) \mu_\varphi(dx), \quad \text{for bounded continuous } f.$$

Theorem 1.3 (LSD as the root spectral measure of a limiting operator). *For any Lévy-Khintchine ensemble $(\mathcal{C}_n)_{n \in \mathbb{N}}$ with characteristics $(\sigma^2, 0, \Pi)$, the limiting spectral distribution $\mu_{\mathcal{C}_\infty}$ of Theorem 1.1 is the expected spectral measure, at the root vector e_\emptyset , of a self-adjoint random conductance operator \mathcal{C}_{G_∞} on $L^2(\mathbb{N}_0^F)$ where G_∞ is a PWIST(σ, λ_Π).*

Remarks:

1. The symmetry of the measure $\mu_{\mathcal{C}_\infty}$ is now easy to see, since every PWIST(σ, λ_Π) is a tree, and thus the odd moments of $\mu_{\mathcal{C}_\infty}$ vanish.
2. It has been pointed out to use that if $(\mathcal{C}_n)_{n \in \mathbb{N}}$ and $(\mathcal{C}'_n)_{n \in \mathbb{N}}$ are independent with characteristics $(0, 0, \Pi)$ and $(\sigma^2, 0, 0)$, then by the Lévy-Itô decomposition, the sequence $(\mathcal{C}_n + \mathcal{C}'_n)_{n \in \mathbb{N}}$ has characteristics $(\sigma^2, 0, \Pi)$. According to Voiculescu's asymptotic freeness theorem, $(\mathcal{C}_n)_{n \in \mathbb{N}}$ and $(\mathcal{C}'_n)_{n \in \mathbb{N}}$ are also asymptotically free, thus one expects that the PWIST($1, \lambda_\Pi$) has an interpretation as a free product of the PWIT(λ_Π) and \mathbb{N} (See also the two remarks following Proposition 1.4 and the first paragraph of Section 5).

Indeed, if \mathcal{A} and \mathcal{B} are two deterministic adjacency operators, then there is a *free product of graphs* associated to their unital C^* -algebra free product $\mathcal{A} * \mathcal{B}$ [Gut98, ALS07]. In a future work, we develop a *free product for random graphs* and show that the PWIST($1, \lambda_\Pi$) has an alternate interpretation as a free product of the PWIT(λ_Π) and \mathbb{N} . See also [Mal11] for some other generalizations of free products which also includes graph free products.

The following result is an application of the resolvent identity, and it may be used in conjunction with Theorem 1.3 to further analyze $\mu_{\mathcal{C}_\infty}$. It can be viewed as an operator version of the Schur complement formula.

Proposition 1.4 (Recursive distributional equation). *Suppose G_∞ is a PWIST(σ, λ_Π). For all $z \in \mathbb{C}_+$ the random variable*

$$R_{\emptyset\emptyset}(z) := \langle e_\emptyset, (\mathcal{C}_{G_\infty} - zI)^{-1} e_\emptyset \rangle$$

satisfies $R_{\emptyset\emptyset}(-\bar{z}) = -\bar{R}_{\emptyset\emptyset}(z)$ and the recursive distributional equation (RDE)

$$R_{\emptyset\emptyset}(z) \stackrel{d}{=} - \left(z + \sigma^2 R_{00}(z) + \sum_{k \in \mathbb{N}} |\mathcal{C}(k)|^2 R_{kk}(z) \right)^{-1} \quad (14)$$

where for all $k \geq 0$, R_{kk} has the same distribution as $R_{\emptyset\emptyset}$ and $\{\mathcal{C}(k)\}_{k \in \mathbb{N}}$ are the points of an independent Poisson random measure with intensity $\Pi(dx)$ on $\mathbb{R} \setminus \{0\}$.

Remarks:

1. One can extend the proposition to non-Hermitian Lévy-Khintchine ensembles using Lemma 2.5 in [BCC11b].

2. For an example of how the above proposition maybe be used, consider Wigner matrices with i.i.d. entries with possibly infinite second moments, but normalized to be in the Gaussian domain of attraction. In this case, the Lévy measure Π is trivial and the $\text{PWIST}(\sigma, 0)$ is just \mathbb{N} rooted at 1.



When there is no Levy measure, the PWIST is the half-line \mathbb{N} . It is well-known that the spectral measure at the root is semi-circle.

Since the edge-weights of the limiting graph are nonrandom, a simple argument shows (see Eq. (41) below) that the resulting recursive equation is the Cauchy-Stieltjes transform (see (40)) of Wigner's semi-circle law:

$$R_{\emptyset\emptyset}(z) = S_{\mu_{sc}}(z) = -\left(z + \sigma^2 S_{\mu_{sc}}(z)\right)^{-1}.$$

3. The formula in (14) is related to subordination formulas from free probability. Let μ_1 and μ_2 denote the LSDs for the characteristics $(0, 0, \Pi)$ and $(\sigma^2, 0, \Pi)$, respectively, and let S_{μ_1} and S_{μ_2} be their Cauchy-Stieltjes transforms. It was shown in [Bia97] that

$$S_{\mu_2}(z) = S_{\mu_1}(z + \sigma^2 S_{\mu_2}(z)). \quad (15)$$

On the other hand, letting $R^{(1)}$ and $R^{(2)}$ denote the resolvent operator for the limiting random conductance operators associated to $(0, 0, \Pi)$ and $(\sigma^2, 0, \Pi)$, respectively, we have that (14) applied to $R^{(1)}(z')$ for $z' := z + \sigma^2 R_{\emptyset\emptyset}^{(2)}$ gives

$$R_{\emptyset\emptyset}^{(1)}(z + \sigma^2 R_{\emptyset\emptyset}^{(2)}) \stackrel{d}{=} -\left(z + \sigma^2 R_{\emptyset\emptyset}^{(2)} + \sum_{k \in \mathbb{N}} |\mathcal{C}(k)|^2 R_{kk}^{(1)}(z + \sigma^2 R_{\emptyset\emptyset}^{(2)})\right)^{-1}$$

which can be compared to (14) applied to $R^{(2)}(z)$.

The rest of the paper is organized as follows. In the next section, we introduce a replacement procedure which creates a new sequence of matrices by modifying a given Lévy-Khintchine ensemble. This modification replaces complex values with real values and also embodies our notion of “cords to infinity”. It is the key procedure which allows us to generalize PWITs to PWISTs. In Section 3, the moment method is used to prove a weak version of Theorem 1.1 in the case that the Lévy measure Π has bounded support. The main point of Section 3, however, is to show that the limiting root spectral measure of a Lévy-Khintchine ensemble is invariant under the replacement procedure of Section 2 (in preparation for proofs of the main results). In Section 4, we precisely define local weak convergence and present an adaptation of the arguments of [BCC11a]. In particular, we show that the local weak convergence argument proves Theorem 1.3 for real Lévy-Khintchine ensembles with $\sigma = 0$. Finally, in Section 5, we combine the arguments of Sections 3 and 4 to prove the main results in the general case. In the appendix we gather some known results which are needed along the way.

Acknowledgments This project arose out of several discussions with Shannon Starr. We thank him for being a sounding board and for his many helpful discussions, insights, and suggestions.

2 A replacement procedure for cords to infinity

In this section, we define an important sequence of *modified* matrices $(\mathcal{C}_n^\sigma)_{n \in \mathbb{N}}$ which play a key role in the proofs of the main results. In particular, these matrices are modifications of a Lévy-Khintchine ensemble $(\mathcal{C}_n)_{n \in \mathbb{N}}$ under a certain replacement procedure which we describe below.

For $h > 0$, by (4) and Proposition A.1 we have that as $n \rightarrow \infty$,

$$\sum_{k=1}^n \pm |\mathcal{C}_n(1, k)| 1_{\{|\mathcal{C}_n(1, k)| \leq h\}}$$

converges in distribution to $ID(\sigma_h^2, 0, \Pi_h)$ where the \pm signs are chosen using independent Rademacher variables (independent also from \mathcal{C}_n), and

$$\begin{aligned} \sigma_h^2 &:= \sigma^2 + \int_{|x| \leq h} x^2 \Pi(dx) \quad \text{and} \\ \Pi_h(dx) &:= 1_{[-h, h]}(x) \Pi(dx). \end{aligned}$$

By a diagonalization argument, we may choose a sequence of positive numbers $h_n \rightarrow 0$ such that we get the following weak convergence to a Gaussian:

$$\sum_{k=1}^n \pm |\mathcal{C}_n(1, k)| 1_{\{|\mathcal{C}_n(1, k)| \leq h_n\}} \Rightarrow \mathcal{N}(0, \sigma^2).$$

In particular, as $h_n \rightarrow 0$

$$\begin{aligned} &\lim_{n \rightarrow \infty} \sum_{k=2}^n \mathbf{E} (|\mathcal{C}_n(1, k)|^2 1_{\{|\mathcal{C}_n(1, k)| \leq h_n\}}) \\ &= \lim_{n \rightarrow \infty} n \mathbf{E} (|\mathcal{C}_n(1, 2)|^2 1_{\{|\mathcal{C}_n(1, 2)| \leq h_n\}}) = \sigma^2. \end{aligned} \tag{16}$$

Our replacement procedure is as follows. For all entries such that $|\mathcal{C}_n(j, k)| > h_n$ as well as for all diagonal entries $\mathcal{C}_n(j, j)$, we set $\mathcal{C}_n^\sigma(j, k) := \pm |\mathcal{C}_n(j, k)|$ where the signs \pm are given by independent Rademacher variables on the upper triangle, and determined on the lower triangle to preserve self-adjointness. However, the entries in positions (j, k) , $j \neq k$ in the matrix \mathcal{C}_n^σ , which satisfy the condition $|\mathcal{C}_n(j, k)| \leq h_n$, will remain blank for now and will be assigned values that are either 0 or σ .

We next describe how to fill in blank entries. We first need to determine the order of the rows (and columns to preserve self-adjointness) by which we fill in the blanks. Recall that \mathcal{C}_n determines a geometric graph, rooted at 1, with edge-weights given by $1/\mathcal{C}_n(j, k)$ as in (12). Let α be the permutation of $\{1, \dots, n\}$ such that $\alpha(i)$ is the i th closest vertex from the root 1 using the distance (35). If j and k are at equal distance from the root 1, we break ties by deeming j “closer” to the root whenever $j < k$. We now fill in blank entries according to the order determined by the (random) permutation α . For instance, we fill in blanks in row 1 first since $\alpha(1) = 1$ (the root is always closest to itself). Next we fill in blank entries in the row $\alpha(2)$, then row $\alpha(3)$, etc.

The procedure for filling in blank entries in row $j = \alpha(i)$ is as follows, starting with row $1 = \alpha(1)$. Out of all k satisfying

$$|\mathcal{C}_n(j, k)| \leq h_n, \quad k \neq j \tag{17}$$

choose one uniformly at random and set this entry, in \mathcal{C}_n^σ , to σ . Set other blank entries in row j , satisfying (17), to zero in the matrix \mathcal{C}_n^σ . This completes the filling of row j of \mathcal{C}_n^σ , and we use the symmetry condition $\mathcal{C}_n^\sigma(j, k) = \mathcal{C}_n^\sigma(k, j)$, to fill in blank entries in the column j .

When row and column $j = \alpha(i)$ are completely filled, we repeat the procedure on row and column $\alpha(i + 1)$. We continue the replacement procedure described in the previous paragraph until all blank entries have been filled, then we say call $(\mathcal{C}_n^\sigma)_{n \in \mathbb{N}}$ the **modified sequence of matrices**.

3 The moment method

In this section, we use the moment method to prove a *convergence in expectation*³ version of Theorem 1.1 in the case where there exists an almost sure bound $0 < \tau < \infty$ on the entries of the Lévy-Khintchine ensemble $(C_n)_{n \in \mathbb{N}}$,

$$|C_n(1, 2)| \leq \tau \quad \text{for all } n. \quad (18)$$

In particular, using the associated Poisson approximation for the distribution of $C_n(1, 2)$ (see [Kal02, Cor. 15.16]) one sees that Π must be supported on $[-\tau, \tau]$.

Let

$$M_p(\mu) := \int_{\mathbb{R}} x^p \mu(dx)$$

be the p th moment of the measure μ . The moment method in this section consists of showing

$$\lim_{n \rightarrow \infty} M_p(\mathbf{E}\mu_{C_n}) = M_p(\mathbf{E}\mu_{C_\infty}), \quad \text{for all } p \in \mathbb{N}, \quad (19)$$

and then verifying that the moments $M_p(\mathbf{E}\mu_{C_\infty})$ determine $\mathbf{E}\mu_{C_\infty}$. However, the main result of this section is the following important consequence of such a verification.

Proposition 3.1 (Invariance of expected LSD under replacement procedure). *If the expected LSD for a Lévy-Khintchine ensemble $(C_n)_{n \in \mathbb{N}}$ exists and is determined by its moments, then it is equal to the limiting expected spectral measure associated to e_1 (the first vector of the standard basis) for any modified sequence $(C_n^\sigma)_{n \in \mathbb{N}}$.*

Proof. A standard argument (see [AGZ10, Ch. 2] or [Tao12, Sec. 2.3.4] for details) shows that the p -moments are given by

$$\begin{aligned} M_p(\mathbf{E}\mu_{C_n}) &= \mathbf{E} \frac{1}{n} \text{tr}(C_n^p) \\ &= \sum_{j_2, \dots, j_p=1}^n \mathbf{E}(C_n(1, j_2)C_n(j_2, j_3) \cdots C_n(j_p, 1)) \end{aligned} \quad (20)$$

where we have set $j_1 = 1$ by exchangeability. The ordered listings of subscript pairs

$$((1, j_2)(j_2, j_3), \dots, (j_p, 1))_{j_2, \dots, j_p=1}^n$$

are viewed as distinct paths of length p which start and end at 1 in the complete graph on $\{1, \dots, n\}$, with edges having orientations, and with the possibility that edges are crossed multiple times. These paths are called *cycles rooted at 1*.

We now make some preliminary observations in order to rewrite (20) as (26). The expression of the p th moment in (26) below allows us to then prove the result.

By Proposition 4 in [Zak06], in the limit as $n \rightarrow \infty$, the only cycles that *contribute to the limiting sum* on the right-side of (20) are “trees” in the following sense. For a given *contributing* term, if the oriented edge (j_k, j_{k+1}) is crossed $q = q(k)$ times, then it must also be crossed q times in the opposite orientation. Thus, for each k there is a corresponding $k' \neq k$ such that

$$C_n(j_k, j_{k+1}) = \overline{C_n(j_{k'}, j_{k'+1})}, \quad j_k = j_{k'+1}, \quad j_{k+1} = j_{k'}. \quad (21)$$

Moreover, the partition of $\{1, \dots, p\}$ which pairs each k with its corresponding k' must be a *non-crossing pair partition* (see [NS06] for details). In particular, p must be even in order to have a non-trivial moment.

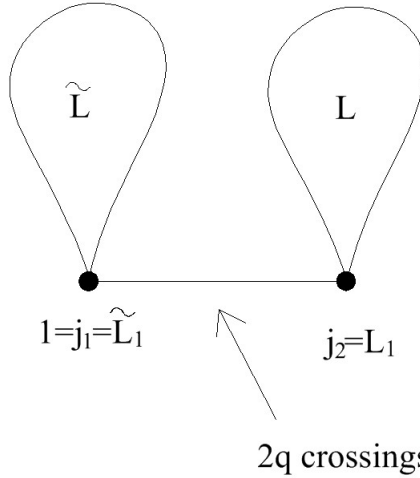
³See [Tao12, Remark 2.4.1] for a definition and short discussion of this type of convergence.

If $\mathcal{C}_n(j_k, j_{k+1})$ appears $q = q(k)$ distinct times in a given term, then its conjugate (or reversed edge from $j_{k'}$ to $j_{k'+1}$) also appears $q = q(k')$ distinct times. Using independence and exchangeability, each term of the sum in (20) takes the form

$$\mathbf{E}|\mathcal{C}_n(1, 2)|^{2q_1} \mathbf{E}|\mathcal{C}_n(1, 2)|^{2q_2} \cdots \mathbf{E}|\mathcal{C}_n(1, 2)|^{2q_\ell} \quad (22)$$

where $2q_1 + \cdots + 2q_\ell = p$.

Fix the value of j_2 and consider a cycle rooted at 1 corresponding to a term in the sum (20) such that $(1, j_2)$ is crossed $q = q(1)$ times in each direction for a total of $2q$ times. Removing these $2q$ edges from our cycle leaves us with several sub-cycles. These sub-cycles can be permuted and then concatenated to form two sub-cycles L and \tilde{L} rooted at $L_1 := j_2$ and $\tilde{L}_1 := 1$ which avoid the edges $(1, j_2)$ and $(j_2, 1)$ (one or both of the cycles may be trivial).



Write $\mathbb{L}(j_2, q)$ for the set of all pairs of cycles (L, \tilde{L}) which are possible, where in particular, different permutations/concatenations leading to the same L or \tilde{L} are each listed separately in $\mathbb{L}(j_2, q)$, i.e., L and \tilde{L} remember their original sub-cycle structure. Also, let s, \tilde{s} be the lengths of L, \tilde{L} so that $s + \tilde{s} = p - 2q$, and write $L \equiv ((L_1, L_2), \dots, (L_s, L_1))$ and similarly for \tilde{L} . Discarding some terms which do not contribute to the limiting sum, we have that (20) can be rewritten as

$$\begin{aligned} & \sum_{q=1}^{p/2} \sum_{j_2=2}^n \sum_{(L, \tilde{L}) \in \mathbb{L}(j_2, q)} \mathbf{E}|\mathcal{C}_n(1, j_2)|^{2q} \mathbf{E}(\mathcal{C}_n(L_1, L_2) \cdots \mathcal{C}_n(L_s, L_1) \mathcal{C}_n(\tilde{L}_1, \tilde{L}_2) \cdots \mathcal{C}_n(\tilde{L}_{\tilde{s}}, \tilde{L}_1)) \\ &= \sum_{q=1}^{p/2} \sum_{j_2=2}^n \left\{ \mathbf{E} \left[|\mathcal{C}_n(1, j_2)|^{2q} (1_{\{|\mathcal{C}_n(1, j_2)| \leq h_n\}} + 1_{\{|\mathcal{C}_n(1, j_2)| > h_n\}}) \right] \times \right. \\ & \quad \left. \sum_{L, \tilde{L} \in \mathbb{L}(j_2, q)} \mathbf{E} \left(\mathcal{C}_n(L_1, L_2) \cdots \mathcal{C}_n(L_s, L_1) \mathcal{C}_n(\tilde{L}_1, \tilde{L}_2) \cdots \mathcal{C}_n(\tilde{L}_{\tilde{s}}, \tilde{L}_1) \right) \right\}. \end{aligned} \quad (23)$$

Recall from (16) that for $\epsilon > 0$, we may find N such that $n \geq N$ implies

$$n \mathbf{E} \left(|\mathcal{C}_n(1, 2)|^2 1_{\{|\mathcal{C}_n(1, 2)| \leq h_n\}} \right) \leq \sigma^2 + \epsilon. \quad (24)$$

which in turn implies

$$n \mathbf{E} \left(|\mathcal{C}_n(1, 2)|^{2q} 1_{\{|\mathcal{C}_n(1, 2)| \leq h_n\}} \right) \leq h_n^{2q-2} (\sigma^2 + \epsilon). \quad (25)$$

To see this, note that a distribution satisfying (24) with maximum $2q$ th moment is given by $\mathcal{C}_n(1, 2) = \pm h_n$ with probability $\frac{\sigma^2 + \epsilon}{nh_n^2}$ and $\mathcal{C}_n(1, 2) = 0$ otherwise. Since $h_n \rightarrow 0$ we see that (25) goes

to zero for $q > 1$. Multiplying out the right side of (23), we have that any term with a factor of $1_{\{|c_n(1,j_2)| \leq h_n\}}$ must have $q(1) = 1$ in order to contribute to the limiting sum. It should perhaps be noted that since we must have that $q(1) = 1$, for terms with a factor of $1_{\{|c_n(1,j_2)| \leq h_n\}}$, the permuting/concatenating of sub-cycles which form L and \tilde{L} is not needed.

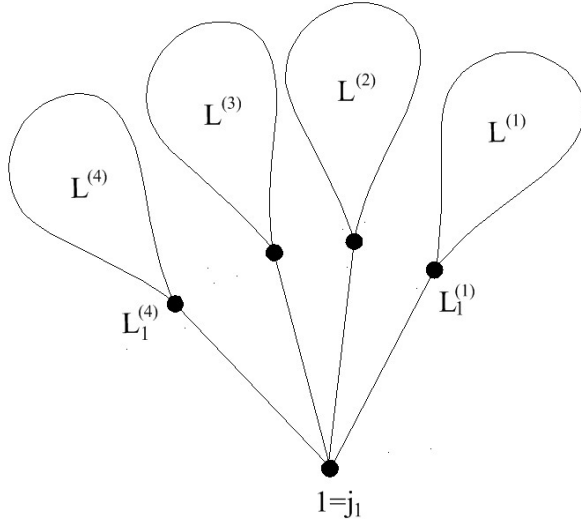
We now write

$$C_n(j_k, j_{k+1}) = C_n(j_k, j_{k+1})(1_{\{|c_n(j_k, j_{k+1})| \leq h_n\}} + 1_{\{|c_n(j_k, j_{k+1})| > h_n\}})$$

for all factors in all terms of (20) and (23). For fixed $j_2 \equiv L_1$, we will categorize terms containing the factor $1_{\{|c_n(1, j_2)| \leq h_n\}}$ by the number of other factors in the term which are of the form

$$|C_n(1, j_k)|^2 1_{\{|c_n(1, j_k)| \leq h_n\}} \quad \text{for any } k.$$

There are at most $p/2$ such factors. In particular, consider terms of (23) which include the factor $|C_n(1, \tilde{L}_2)|^2 1_{\{|c_n(1, \tilde{L}_2)| \leq h_n\}}$. The above procedure on our cycle rooted at 1 is repeated on the cycle $\tilde{L} =: \tilde{L}^{(1)}$, which is also rooted at 1. In other words, we fix the value of \tilde{L}_2 and consider cycles such that the edge $(1, \tilde{L}_2)$ is crossed exactly once in each direction. We remove these 2 edges from \tilde{L} leaving us with two sub-cycles $L^{(2)}$ and $\tilde{L}^{(2)}$ rooted at $L_1^{(2)} := \tilde{L}_2^{(1)}$ and $\tilde{L}_1^{(2)} := 1$. We then repeat the procedure on the cycle $\tilde{L}^{(2)}$ to get two more sub-cycles $L^{(3)}$ and $\tilde{L}^{(3)}$, and continue this process until all edges of the form $(1, \cdot)$ or $(\cdot, 1)$ are “removed”. Thus, for any term containing $1_{\{|c_n(1, j_2)| \leq h_n\}}$ there is a corresponding list of cycles $(L^{(1)}, L^{(2)}, \dots, L^{(M)})$. The list is of length $M \leq p/2$ where M depends on the term (thus terms are categorized by their associated M value), and each cycle in the list is rooted at a different vertex in $\{2, \dots, n\}$. Let $\mathbb{L}_M(n)$ denote the set of all possible lists of cycles of length M .



Finally, recalling that $L_1^{(1)} \equiv j_2$, the sum of all contributing terms in (23) can be written in the form

$$\sum_{M=0}^{p/2} \sum_{\mathbb{L}_M(n)} \prod_{i=1}^M \left(\mathbf{E} \left[|C_n(1, L_1^{(i)})|^2 1_{\{|c_n(1, j_2)| \leq h_n\}} \right] \mathbf{E} \left[\prod_{i=1}^M C_n(L_1^{(i)}, L_2^{(i)}) \cdots C_n(L_{s(i)}^{(i)}, L_1^{(i)}) \right] \right).$$

Summing over the possible first coordinates of each cycle in the list of cycles, $L_1^{(i)} \in \{2, \dots, n\}$,

and taking the limit gives us

$$\lim_{n \rightarrow \infty} \sum_{M=0}^{p/2} \sum_{(L^{(1)}, \dots, L^{(M)}) \in \mathbb{L}_M(n)} \sigma^{2M} \mathbf{E} \left[\prod_{i=1}^M \mathcal{C}_n(L_1^{(i)}, L_2^{(i)}) \cdots \mathcal{C}_n(L_{s^{(i)}}^{(i)}, L_1^{(i)}) \right]. \quad (26)$$

Recall that $(\mathcal{C}_n^{\sigma,1})$ are matrices which are modified using only the first step of the replacement procedure, i.e., where only a single cord to infinity (from 1) has been substituted. Using the fact that

$$|\mathcal{C}_n(j_k, j_{k+1})| = |\mathcal{C}_n^{\sigma,1}(j_k, j_{k+1})| \text{ on the event } \{|\mathcal{C}_n(j_k, j_{k+1})| > h_n\},$$

a relatively straightforward calculation of $M_p(\mathbf{E}\mu_{\tilde{\mathcal{C}}_n^\sigma})$ using (20) also gives (26) by realizing that

- (a) the number of times that a given cycle rooted at 1 crosses the cord from 1 to infinity in either direction is exactly $2M$, and
- (b) for a fixed set of loops $L^{(1)}, \dots, L^{(M)}$ with different roots, one can identify the different roots with one single root. One need only check that the two configurations of loops give the same value for the expression

$$\lim_{n \rightarrow \infty} \mathbf{E} \left[\prod_{i=1}^M \mathcal{C}_n(L_1^{(i)}, L_2^{(i)}) \cdots \mathcal{C}_n(L_{s^{(i)}}^{(i)}, L_1^{(i)}) \right]. \quad (27)$$

There is a slight subtlety regarding the invariance of (27) under the identification of roots. First of all, we can approximate the Lévy measure by a sum of Dirac point measures, and without loss of generality, we thus assume it has this form. The subtlety is that the dependence structure of edges crossed in a given path is changed under the identification of roots. But it turns out that the dependence structure does not affect (27) since (i) the dependence structure only changes on the event that edges have a common weight λ , and (ii) the $2q$ th moment of λ times a Rademacher random variable is λ^{2q} . Thus, for example, the product of the variances of two independent λ -scaled Rademachers is exactly the fourth moment of a single λ -scaled Rademacher.

The proof of the theorem is now complete for the first step of the replacement procedure. Equivalence of moments for other steps in the replacement procedure follows similarly, and the rest of the proof is left as an exercise. \square

Remark. When Π is trivial, all the q_i 's in (22) are *all* equal to 2. This leads to the well-known fact that (20) is the number of Dyck words of length $2p$ which is just the p th Catalan number

$$c_p = \frac{(2p)!}{p!(p+1)!}.$$

We next have a result which relates the moments of the matrix entries to the moments of the Lévy measure. Both sets of moments are also related to the moments of the LSD using (22); moreover, together with the proposition below, (22) proves existence of the limit in (19).

Proposition 3.2 (Triangular array moments are related to Lévy measure moments). *Suppose that $\{\mathcal{C}(n, k), 1 \leq k \leq n\}_{n \in \mathbb{N}}$ is a triangular array of random variables which are i.i.d. in each row, and for which $\sum_{k=1}^n |\mathcal{C}(n, k)|^2$ converges weakly as $n \rightarrow \infty$ to an infinitely divisible law with subordinator characteristics (σ^2, Π_s) . If the random variables are uniformly bounded,*

$$|\mathcal{C}(n, k)| \leq \tau \text{ for all } n \text{ and } k, \quad (28)$$

then

$$\lim_{n \rightarrow \infty} n \mathbf{E} |\mathcal{C}(n, 1)|^2 = \sigma^2 + M_1(\Pi_s),$$

and for $p > 1$

$$\lim_{n \rightarrow \infty} n \mathbf{E} |\mathcal{C}(n, 1)|^{2p} = M_p(\Pi_s).$$

Proof. Set $X_n := |\mathcal{C}(n, 1)|^2$ with characteristic function φ_{X_n} . The characteristic function of

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n |\mathcal{C}(n, 1)|^2 \stackrel{d}{=} X$$

in (3) takes the form

$$\varphi_X(\theta) = \exp \left(i\theta\sigma^2 + \int_0^{\tau^2} (e^{i\theta x} - 1) \Pi_s(dx) \right) \quad (29)$$

and by convergence in distribution of the row sums and Lemma 5.8 in [Kal02],

$$\lim_{n \rightarrow \infty} n(\varphi_{X_n} - 1) = i\theta\sigma^2 + \int_0^{\tau^2} (e^{i\theta x} - 1) \Pi_s(dx)$$

uniformly in θ on compact subsets of \mathbb{R} . Since the X_n are bounded and since Π_s has bounded support we may expand both sides in terms of power series and switch summations with integrals. This gives us

$$\lim_{n \rightarrow \infty} n \sum_{k \geq 1} \frac{(i\theta)^k \mathbf{E} X_n^k}{k!} = i\theta\sigma^2 + \sum_{k \geq 1} \int_0^{\tau^2} \frac{(i\theta x)^k}{k!} \Pi(dx) \quad (30)$$

uniformly on compact subsets, from which the lemma follows. \square

To verify the “moment problem” required to use Proposition 3.1, we adapt arguments from [BG01, KSV04, Zak06]. Let Q_p be the set of (q_1, \dots, q_ℓ) such that $q_i \in \mathbb{N}$, $\sum_{i=1}^\ell q_i = p$, and

$$q_1 \geq q_2 \geq \dots \geq q_\ell.$$

Also, fix a sequence of distinct colors $\{K_i\}_{i=0}^\infty$. We define $T((q_1, \dots, q_p))$ to be the number of colored rooted trees which satisfy

- There are $p + 1$ vertices.
- There are exactly q_i vertices of color K_i with the root being the only vertex of color K_0 .
- If u and v are the same color then the distance from u to the root is equal to the distance from v to the root.
- If u and v have the same color then so do their parents.

Define

$$\mathcal{I}_{p,\ell} := \sum_{(q_1, \dots, q_\ell) \in Q_p} T((q_1, \dots, q_\ell)).$$

Proposition 3.3 (LSD determined by its moments). *Under assumption (18),*

$$M_{2p}(\mathbf{E}\mu_{\mathcal{C}_\infty}) \leq \tau^{2p} \sum_{\ell} \mathcal{I}_{p,\ell} (M_2(\Pi) + \Pi([-1, 1]^c) + \sigma^2)^\ell, \quad (31)$$

and thus $\mathbf{E}\mu_{\mathcal{C}_\infty}$ exists and is determined by its moments.

Proof. By splitting the support of Π into $[-1, 1]$ and its complement, note that $M_{2q}(\Pi) \leq M_2(\Pi) + \tau^{2q}\Pi([-1, 1]^c)$. Also, without loss of generality, $\tau \geq 1$. We use Lemma 3.2 in conjunction with the argument of [Zak06, Thm. 2] (see also [BG01, Sec. 5.3] and [KSV04, Sec. IV]) to get

$$\begin{aligned}
M_{2p}(\mathbf{E}\mu_{\mathcal{C}_\infty}) &= \\
&\lim_{n \rightarrow \infty} \sum_{(q_1, \dots, q_\ell) \in \mathcal{Q}_p} T((q_1, \dots, q_\ell)) n \mathbf{E}(|\mathcal{C}_n(1, 2)|^{2q_1}) \cdots n \mathbf{E}(|\mathcal{C}_n(1, 2)|^{2q_\ell}) \\
&\leq \sum_{(q_1, \dots, q_\ell) \in \mathcal{Q}_p} T((q_1, \dots, q_\ell)) (M_{2q_1}(\Pi) + \sigma^2) \cdots (M_{2q_\ell}(\Pi) + \sigma^2) \\
&\leq \tau^{2p} \sum_{(q_1, \dots, q_\ell) \in \mathcal{Q}_p} T((q_1, \dots, q_\ell)) (M_2(\Pi) + \Pi([-1, 1]^c) + \sigma^2)^\ell \\
&= \tau^{2p} \sum_{\ell} \mathcal{I}_{p, \ell} (M_2(\Pi) + \Pi([-1, 1]^c) + \sigma^2)^\ell.
\end{aligned} \tag{32}$$

Next, we use Eq. (9) in [BG01] which gives the bound

$$\mathcal{I}_{p, \ell} \leq c_p \mathcal{S}_{p, \ell} \tag{33}$$

where c_p is the p th Catalan number and

$$\mathcal{S}_{p, \ell} = \frac{1}{\ell!} \sum_{k=0}^{\ell} (-1)^{\ell-k} \binom{\ell}{k} k^{2p}$$

is a Stirling number of the second kind. By (32), (33), and Theorem 30.1 in [Bil86], $\mathbf{E}\mu_{\mathcal{C}_\infty}$ is determined by its moments if for any $R > 0$,

$$\frac{c_p}{(2p)!} \sum_{\ell=1}^p R^\ell \mathcal{S}_{p, \ell} \tag{34}$$

is $o(r^p)$ for some r as $p \rightarrow \infty$, and this is easily verified. For example Section 5.5 of [BG01] shows (34) is less than $(p^p + e^{R(p-1)})/(p!(p+1)!)$. \square

Remark. In [BG01], the lower bound $\mathcal{S}_{2p, \ell} \leq \mathcal{I}_{2p, \ell}$ was also established and used to show that the LSD has unbounded support (see also [Zak06, Prop. 12]). In our situation, this tells us that the Lévy-Khintchine ensembles for which the LSD has bounded support are precisely those with only a Wigner portion, i.e., those with characteristics of the form $(\sigma^2, 0, 0)$.

4 From local weak convergence to spectral convergence

In this section, to simplify things we restrict our attention to random conductance matrices \mathcal{C}_n with *real* entries. The goal of this section is to present Theorem 4.2 which uses strong resolvent convergence to connect the notions of local weak convergence and weak convergence of ESDs. Theorem 4.2 below is similar to [BCC11a, Theorem 2.2] (see also [BL10, BCC11b, BC12]), and its proof is an adaptation of the arguments there which treat the symmetric α -stable case:

$$(\sigma^2, 0, \Pi) = (0, 0, \text{sign}(x)\alpha|x|^{-1-\alpha}dx).$$

Here we replace the α -stable Lévy measure with an arbitrary symmetric Lévy measure $\Pi(dx)$ on $\mathbb{R} \setminus \{0\}$. In particular, if one assumes self-adjointness of the limiting operator (which follows from

Lemma 5.2 below), then the arguments in this section are enough to handle Theorem 1.3 in the case when $\sigma = 0$ and the entries are real.

Let us now present the precise notion of local weak convergence following the treatment in [AS04]. Let $G[\emptyset] = (V, E)$ be a \emptyset -rooted graph with vertex set V and edge set E both of which are at most countably infinite. Any edge-weight function $\mathcal{R} : E \rightarrow \mathbb{R} \setminus \{0\}$ defines a distance between any two vertices $u, v \in V$ as

$$d(u, v) := \inf_{\gamma \text{ connects } u, v} \sum_{e \in \gamma} |\mathcal{R}(e)| \quad (35)$$

where the infimum is over all paths γ which connect vertices u and v . The distance d naturally turns $G[\emptyset]$ into a metric space. We include $\pm\infty$ as a possible edge-weight where $\pm\infty$ is thought of as the same weight using the one-point compactification of $\mathbb{R} \setminus \{0\}$.

If $G[\emptyset]$ is connected and undirected and the edge-weight function \mathcal{R} is such that for every vertex v and every $r < \infty$, the number of vertices within distance r of v is finite, then $G[\emptyset] = (V, E, \mathcal{R})$ is a **rooted geometric graph**. Henceforth all graphs will be rooted geometric graphs, and when they are rooted at the default root \emptyset , we may simply write G instead of $G[\emptyset]$. The set of all rooted geometric graphs is written \mathcal{G}_* .

In the case that the range of \mathcal{R} is positive and the underlying graph is a tree, we can interpret \mathcal{R} as assigning resistances to edges. However, for technical reasons required by the proofs of our main results, *we allow \mathcal{R} to take negative values*. The possibility of negative weights makes our treatment here differ slightly from [AS04]. But, using the modulus in (35) nevertheless permits us to reap the benefits of the metric of [AS04] on \mathcal{G}_* .

Let $\mathcal{N}_{r, \emptyset}(G)$ be the r -neighborhood of \emptyset . This is the \emptyset -rooted subgraph of G formed by restricting the graph to the set of all vertices $v \in V$ such that $d(\emptyset, v) \leq r$ and restricting to the set of edges that can be crossed by journeying at most distance r from the root \emptyset . We say r is a *continuity point* of G if there is no vertex of exact distance r from the root.

Definition 4.1 (The topology of \mathcal{G}_*). We say $(G_n = (V_n, E_n, \mathcal{R}_n))_{n \in \mathbb{N}}$ converges to $G = (V, E, \mathcal{R})$ in \mathcal{G}_* if for each continuity point r of G , there is an n_r such that $n > n_r$ implies there exists a graph isomorphism

$$\pi_n : \mathcal{N}_{r, \emptyset}(G) \rightarrow \mathcal{N}_{r, \emptyset}(G_n)$$

which preserves the root and for which

$$\lim_{n \rightarrow \infty} \mathcal{R}_n(\pi_n^{-1}(u), \pi_n^{-1}(v)) = \mathcal{R}(u, v). \quad (36)$$

As noted in [AS04], the above convergence determines a topology which turns \mathcal{G}_* into a complete separable metric space. Using the usual theory of convergence in distribution, one can therefore say that a sequence of random rooted geometric graphs $(G_n)_{n \in \mathbb{N}} \subset \mathcal{G}_*$, with distributions μ_n , converge weakly to $G \in \mathcal{G}_*$ with distribution μ if for all bounded continuous $f : \mathcal{G}_* \rightarrow \mathbb{R}$

$$\int_{\mathcal{G}_*} f d\mu_n \rightarrow \int_{\mathcal{G}_*} f d\mu. \quad (37)$$

Such weak convergence is called *local weak convergence*.

The following connection between local weak convergence and strong resolvent convergence was first noticed in [BL10] and [BCC11a] in the context of sparse matrices and heavy-tailed matrices, respectively (see [HO07] for related arguments).

Theorem 4.2 (Local weak convergence implies strong resolvent convergence). *Let $(G_n)_{n \in \mathbb{N}}$, which are associated to $(G_n = (V_n, E_n, \mathcal{R}_n))_{n \in \mathbb{N}}$ as in (13), be essentially self-adjoint. Suppose that the*

graphs converge in the local weak sense to a tree $G = (V, E, \mathcal{R})$ with respect to the isomorphisms π_n , and that \mathcal{C}_G is also essentially self-adjoint.

If for each $u \in V$,

$$\lim_{\epsilon \searrow 0} \lim_{n \rightarrow \infty} \sum_{v \in V_n: v \sim \pi_n(u)} |\mathcal{C}_{G_n}(\pi_n(u), v)|^2 1_{\{|\mathcal{C}_{G_n}(\pi_n(u), v)|^2 \leq \epsilon\}} = 0 \text{ a.s.}, \quad (38)$$

then for all $z \in \mathbb{C}_+$, as $n \rightarrow \infty$:

$$\langle e_\emptyset, (\mathcal{C}_{G_n} - zI)^{-1} e_\emptyset \rangle \xrightarrow{w} \langle e_\emptyset, (\mathcal{C}_G - zI)^{-1} e_\emptyset \rangle. \quad (39)$$

Remark. By Proposition A.1, condition (38) simply says that $\sigma^2 = 0$ in (5).

Once one checks the local weak convergence of $(G_n[1])_{n \in \mathbb{N}}$ to a PWIT(λ_Π) and verifies self-adjointness, then the above result essentially handles the case where the Wigner component vanishes. Let us briefly outline this. First of all $\sigma = 0$ will imply condition (38). Next, recall that the Cauchy-Stieltjes transform (or simply Stieltjes transform) is defined as

$$S_\mu(z) := \langle \mu, (x - z)^{-1} \rangle = \int_{\mathbb{R}} \frac{\mu(dx)}{x - z}, \quad z \in \mathbb{C} \setminus \mathbb{R}. \quad (40)$$

Recall from (6) that $\mu_{\mathcal{C}_n}$ is the ESD of \mathcal{C}_n . Using the fact that entries in \mathcal{C}_n are i.i.d.,

$$S_{\mathbf{E}\mu_{\mathcal{C}_n}}(z) = \mathbf{E}S_{\mu_{\mathcal{C}_n}}(z) = \frac{1}{n} \mathbf{E} \text{tr}(\mathcal{C}_n - zI)^{-1} = \mathbf{E}(\mathcal{C}_n - zI)^{-1}(1, 1). \quad (41)$$

Therefore, by (41), the above theorem, and a bound on the modulus of the Green's function

$$|(\mathcal{C}_n - zI)^{-1}(1, 1)| \leq (\Im z)^{-1}$$

for $z \in \mathbb{C} \setminus \mathbb{R}$, we obtain convergence of $S_{\mathbf{E}\mu_{\mathcal{C}_n}}$ to $S_{\mathbf{E}\mu_{\mathcal{C}_\infty}}$ where G_∞ is a PWIT(λ_Π). Lemma A.2, which tells us that the Cauchy-Stieltjes transform determines the LSD, then implies weak convergence of the expected ESDs (since e_\emptyset has unit norm, the limit is a probability measure). A concentration of measure argument from [GL09], Lemma 5.1 below, extends this to a.s. weak convergence for the random ESDs.

For the proof of Theorem 4.2 we need a lemma which appears as Thm VIII.25 in [RS80]. We state it without proof.

Lemma 4.3 (Strong resolvent convergence characterization). *Suppose \mathcal{C}_n and \mathcal{C}_∞ are self-adjoint operators on $L^2(V)$ with a common core \mathcal{D} (for all n and ∞). If*

$$\mathcal{C}_n \varphi \rightarrow \mathcal{C}_\infty \varphi \quad \text{in } L^2(V),$$

for each $\varphi \in \mathcal{D}$, then \mathcal{C}_n converges to \mathcal{C}_∞ in the strong resolvent sense.

Proof of Theorem 4.2. To match the setting for which we employ this theorem, let the vertex set of G_n be a subset of \mathbb{N} and the vertex set of G be \mathbb{N}^F . By assumption, the local weak limit of $(G_n)_{n \in \mathbb{N}}$ is the tree G , with respect to the mappings

$$\pi_n : \mathbb{N}^F \rightarrow V_n \subset \mathbb{N} \quad (42)$$

which are injective when restricted to some random subset of \mathbb{N}^F with the same cardinality as V_n . By the Skorokhod representation theorem we will in fact assume that this weak convergence in \mathcal{G}_* is almost sure convergence on some probability space. Note that when the sequence $(\mathcal{C}_{G_n})_{n \in \mathbb{N}}$ is a

sequence of $n \times n$ Lévy-Khintchine matrices, one may set $V_n = \{1, \dots, n\}$, however in general V_n may even be infinite (in which case it is just \mathbb{N}).

Since \mathbb{N}^F is countable we can fix some bijection with \mathbb{N} and think of V_n as a subset of \mathbb{N}^F . In this case, the maps π_n can each be extended to (random) bijections from \mathbb{N}^F to \mathbb{N} , and abusing notation we write π_n for these extensions. The essentially self-adjoint operators \mathcal{C}_{G_n} extend to self-adjoint operators on $L^2(\mathbb{N}^F)$, using the core \mathcal{D}_{fs} consisting of vectors with finite support, by defining

$$\langle e_u, \mathcal{C}_{G_n} e_v \rangle := \begin{cases} \mathcal{C}_{G_n}(\pi_n(u), \pi_n(v)) & \text{if } \{\pi(u), \pi(v)\} \subset V_n \\ 0 & \text{otherwise.} \end{cases} \quad (43)$$

By assumption, the closure of \mathcal{C}_G is also self-adjoint using the core \mathcal{D}_{fs} . Again abusing notation, we identify this closure with \mathcal{C}_G .

By local weak convergence and Skorokhod representation, we have that almost surely

$$\langle e_u, \mathcal{C}_{G_n} e_v \rangle \rightarrow \langle e_u, \mathcal{C}_G e_v \rangle. \quad (44)$$

By Lemma 4.3, we are left to show that

$$\sum_{u \in \mathbb{N}^F} |\langle e_u, \mathcal{C}_{G_n} e_v \rangle - \langle e_u, \mathcal{C}_G e_v \rangle|^2 \rightarrow 0$$

almost surely, as $n \rightarrow \infty$. This follows from the Vitali convergence theorem since (44) provides almost sure convergence and (38) provides uniform square integrability. \square

A common tool for showing local weak convergence is the following lemma about Poisson random measures which is similar to [Ste02, Lemma 4.1].

Lemma 4.4 (Convergence to a Poisson random measure). *Suppose $\{\mathcal{C}(n, k), 1 \leq k \leq n\}_{n \in \mathbb{N}}$ is a triangular array of real random variables which are i.i.d. in each row, and for which $\sum_{k=1}^n \mathcal{C}(n, k)$ converges in law, as $n \rightarrow \infty$, to an $ID(\sigma^2, b, \Pi)$ random variable. Then as $n \rightarrow \infty$*

$$\sum_{k=1}^n \delta_{\mathcal{C}(n, k)}$$

converge vaguely, as measures on $\mathbb{R} \setminus \{0\}$, to a Poisson random measure η with intensity $\mathbf{E}\eta = \Pi$.

Proof of Lemma 4.4. Note that any Lévy measure Π is also a Radon measure on $\mathbb{R} \setminus \{0\}$. Even though there is a possible singularity at 0, this is no concern since $0 \notin \mathbb{R} \setminus \{0\}$. Therefore, by the basic convergence theorem of empirical measures to Poisson random measures (see Theorem 5.3 in [Res07]) we need only check that

$$n\mathbf{P}(\mathcal{C}_n(1, 2) \in \cdot) \xrightarrow{\text{vag}} \Pi$$

vaguely as measures on $\mathbb{R} \setminus \{0\}$. This follows from Proposition A.1. \square

Remark. It is instructive to recognize that the Lévy characteristics σ^2 and b bear no influence on the above lemma, and consequently bear no influence on local weak convergence of the associated graphs. This is because vague convergence pushes any affect they have to the point 0 which is not in $\mathbb{R} \setminus \{0\}$. This essentially tells us that b has no effect on the LSD which is one reason why we were allowed to set it to 0 (this statement is made rigorous by Theorem 1.3). The same is not true for σ^2 since we must have $\sigma = 0$ in order to satisfy (38) (uniform square integrability) and therefore to use Theorem 4.2. However, after one applies the replacement procedure, (38) will once again be satisfied.

The following proposition utilizes Lemma 4.4 to show local weak convergence to a PWIST. It is a variant of results in [Ald92, Sec. 3] (see also [Ald01, Ste02, BCC11a]).

Proposition 4.5 (Local weak convergence to a PWIST). *Let $G_n[1]$ have conductances $\{\mathcal{C}_n^\sigma(j, k)\}_{j,k}$ which are modified Lévy-Khintchine matrices with characteristics $(\sigma^2, 0, \Pi)$ (modified as in Section 2). Then the local weak limit of $(G_n[1])_{n \in \mathbb{N}}$ is a PWIST (σ, λ_Π) .*

Proof. We follow [Ald92, Sec. 3] and [BCC11a, Sec. 2.5]. For each fixed realization of the $\{\mathcal{C}_n^\sigma(j, k), 1 \leq j, k \leq n\}$ we consider their reciprocals, i.e., the resistances

$$\{\mathcal{R}_n^\sigma(j, k), 1 \leq j, k \leq n\}.$$

For any $B, H \in \mathbb{N}$, such that

$$\sum_{\ell=0}^H B^\ell \leq n,$$

we define a rooted geometric subgraph $G_n[1]^{B,H}$ of $G_n[1]$, whose vertex set is in bijection with a B -ary tree of depth H rooted at 1. Let $V_n := \{1, \dots, n\}$. The bijection provides a partial index of vertices of $G_n[1]$ as elements in

$$J_{B,H} = \bigcup_{\ell=0}^H \{1, \dots, B\}^\ell \subset \mathbb{N}_0^f \quad (45)$$

where the indexing is given by an injective map

$$\pi_n : J_{B,H} \rightarrow V_n.$$

The map π_n easily extends to a bijection from some subset of \mathbb{N}_0^f to V_n and thus can be thought of as restrictions of the maps of (42).

We set $I_\emptyset = \{1\}$ and set the preimage/index of the root 1 to be $\pi_n^{-1}(1) = \emptyset$. We next index the B vertices in $V_n \setminus I_\emptyset$ which have the B smallest absolute values among $\{\mathcal{R}_n^\sigma(1, k)\}_{2 \leq k \leq n}$. The k th smallest absolute value is given the index $\emptyset k = \pi_n^{-1}(v)$, $1 \leq k \leq B$. As in the discussion preceding (11), we have written the vector $\emptyset k$ using concatenation. Breaking ties using the lexicographic order, this defines the first generation.

Now let I_1 be the union of I_\emptyset and the B vertices that have been selected. If $H \geq 2$, we repeat the indexing procedure for the vertex indexed by $\emptyset 1$ (the first child of \emptyset) on the set $V_n \setminus I_1$. We obtain a new set $\{11, \dots, 1B\}$ of vertices sorted by their absolute resistances. We define I_2 as the union of I_1 and this new collection. Repeat the procedure for $\emptyset 2$ on $V_n \setminus I_2$ and obtain a new set $\{21, \dots, 2B\}$. Continuing on through $\{B1, \dots, BB\}$, we have constructed the second generation, at depth 2, and we have indexed a total of $(B^3 - 1)/(B - 1)$ vertices. The indexing procedure is repeated through depth H so that $(B^{H+1} - 1)/(B - 1)$ vertices are sorted. Call this set of vertices $V_n^{B,H} = \pi_n(J_{B,H})$. The subgraph of $G_n[1]$ generated by the vertices $V_n^{B,H}$ is denoted $G_n[1]^{B,H}$ (by “generated” we mean that we include only edges with endpoints in the specified vertex set). It is the modification of $G_n[1]$ such that any edge with at least one endpoint in the complement of $V_n^{B,H}$ is given an infinite resistance. In $G_n[1]^{B,H}$, the elements of $\{u1, \dots, uB\}$ are the children of u .

Note that while the vertex set $V_n^{B,H}$ has a natural tree structure, $G_n[1]^{B,H}$ is actually a subgraph of a complete graph which may not be a tree.

Let $G_\infty[\emptyset]$ be a PWIST (σ, λ_Π) , or a PWIT (λ_Π) if $\sigma = 0$, and write $G_\infty[\emptyset]^{B,H}$ for the finite rooted geometric graph obtained by the sorting procedure just described. Namely, $G_\infty[\emptyset]^{B,H}$ consists of the subtree with vertices of the form $u \in J_{B,H}$, with resistances between these vertices

inherited from the infinite tree. If an edge is not present in $G_\infty[\emptyset]^{B,H}$, we may think of it as being present but having infinite resistance.

Since the conductances $\{C_n^\sigma(j, k)\}$ by definition are real with a symmetric distribution, we may without loss of generality replace $\sum_{j=1}^n \pm |C_n(1, j)|$ with $\sum_{j=1}^n C_n(1, j)$ in (4). We use Lemma 4.4 on the unmodified matrices (with real and symmetrically distributed entries) to conclude that $\sum_{k=1}^n \delta_{C_n(1, k)}$ converges vaguely to a Poisson random measure with intensity Π . For $h_n = L(n)/\sqrt{n}$, the truncation $C(n, k)1_{|C(n, k)| \leq h_n}$ does not affect this vague convergence. Note that besides the random resistances on edges given by the Poisson random measure, there is also one more nonrandom resistance given by the replacement procedure (for n large enough), and the value is always $1/\sigma$. It is easily verified that the property in (36) is satisfied by each edge (u, v) of the tree $G_\infty[\emptyset]$.

It remains to check that for each B and H , our maps π_n are graph isomorphisms for n large enough. In other words, we must check that for each edge in $G_\infty[\emptyset]^{B,H}$ with an infinite resistance, the corresponding edges of $(G_n[1]^{B,H})_{n \in \mathbb{N}}$ (for n large enough), must have resistances which diverge to infinity. The divergence of these resistances to infinity follows from a standard coupling argument which shows that these resistances stochastically dominate i.i.d. variables with distribution $\mathcal{R}_n(1, 2)$ which clearly diverges as $n \rightarrow \infty$ (see for example, Lemma 2.7 in [BCC11a]). \square

5 Proofs of the main results

In the case that a Lévy-Khintchine ensemble $(C_n)_{n \in \mathbb{N}}$ has characteristics of the form $(0, 0, \Pi)$, then results of Section 4 (Theorem 4.2, Proposition 4.5) imply the existence of the LSD in expectation. On the other hand, if $|C_n(1, 2)|$ is a.s. uniformly bounded in n , Proposition 3.3 proves the existence of the LSD in expectation. Armed with our local weak convergence arguments, let us begin this section with a somewhat simple proof of existence of the LSD in expectation under the assumption (18). The proof is based on free probability; for an introduction to the subject we refer the reader to [AGZ10, Ch. 5].

Recall that if two self-adjoint random variables x and y in a C^* -probability space are freely independent with distributions μ and ν , respectively, then the distribution of their sum is the free convolution $\mu \boxplus \nu$.

Suppose $(C_n)_{n \in \mathbb{N}}$ and $(C'_n)_{n \in \mathbb{N}}$ are independent Lévy-Khintchine ensembles with characteristics $(\sigma^2, 0, 0)$ and $(0, 0, \Pi)$ and such that the LSD $\mu_{C'_\infty}$ exists and has p th moments bounded by r^p for some $0 < r < \infty$. Then Voiculescu's asymptotic freeness theorem ([AGZ10, Thm 5.4.5]) implies that the LSDs are freely independent and the LSD of $(C_n + C'_n)_{n \in \mathbb{N}}$ exists and is equal to

$$\mu_{C_\infty} \boxplus \mu_{C'_\infty}.$$

Existence of the LSD for the ensemble $(C''_n)_{n \in \mathbb{N}}$ with a.s. uniformly bounded $|C''_n(1, 2)|$ follows by using the Lévy-Itô decomposition to find independent C_n and C'_n such that

$$C''_n = C_n + C'_n. \quad (46)$$

We have that μ_{C_∞} exists by Wigner's theorem while $\mu_{C'_\infty}$ exists by local weak convergence arguments. The moments of the latter LSD are controlled using Proposition 3.3. Thus $\mu_{C''_\infty}$ exists.

We return now to the general assumptions of Theorems 1.1 and 1.3. Before proving the main results, we have three preliminary lemmas. Our first preliminary lemma allows us to extend from convergence in expectation to almost sure convergence. It is a concentration of measure result first noticed in [GL09, Theorem 1] and later in [BCC11b, Lemma C.2]. We state it here without proof.

Lemma 5.1 (Concentration for ESDs). *Let \mathcal{H}_n be an $n \times n$ Hermitian matrix whose rows are independent (as vectors). For every real-valued continuous $f(x)$ going to 0 as $x \rightarrow \pm\infty$ such that*

$\|f\|_{TV} \leq 1$, and for every $t \geq 0$,

$$\mathbf{P} \left(\left| \int_{\mathbb{R}} f d\mu_{\mathcal{H}_n} - \mathbf{E} \int_{\mathbb{R}} f d\mu_{\mathcal{H}_n} \right| \geq t \right) \leq 2 \exp(-nt^2/2)$$

The next lemma verifies the self-adjointness of PWISTs required to use Theorem 4.2.

Lemma 5.2 (Self-adjointness of PWIST operators). *Suppose $G_{\infty}[\emptyset] = (V_{\infty}, E_{\infty}, \mathcal{R}_{\infty})$ is a PWIST(σ, λ_{Π}). Then the associated random conductance operator $\mathcal{C}_{G_{\infty}}$ on $L^2(V_{\infty})$, as defined in (13), is essentially self-adjoint.*

Proof. Denote the children of the root \emptyset of a PWIST(σ, λ_{Π}) by $\mathbb{N}[\emptyset]$. For $\kappa > 0$, define the random variables

$$\tau_{\kappa} := \begin{cases} 0 & \text{if } \sum_{v \in \mathbb{N}[\emptyset]} |\mathcal{C}_{G_{\infty}}(\emptyset, v)|^2 \leq \kappa \\ 1 & \text{otherwise.} \end{cases}$$

For each symmetric Lévy measure Π and $\sigma > 0$, we may choose κ large enough so that

$$\mathbf{P}(\tau_{\kappa} = 0) > 0$$

and consequently $\mathbf{E}\tau_{\kappa} < 1$. We may therefore employ the proof of Proposition A.2 in [BCC11a] to show that for any PWIST, $G_{\infty} = (V_{\infty}, E_{\infty}, \mathcal{R}_{\infty})$, there is a constant $\kappa > 0$ and a sequence of connected finite increasing subsets $(V_n)_{n \in \mathbb{N}}$ whose union is V_{∞} , and such that for all n and $u \in V_n$

$$\sum_{v \notin V_n : v \sim u} |\mathcal{C}_{G_{\infty}}(u, v)|^2 < \kappa.$$

Finally, the existence of such a κ allows us to use Lemma A.3 in [BCC11a] to conclude that any PWIST is essentially self-adjoint. Thus its closure is self-adjoint. \square

The final preliminary lemma, similar to arguments in [BAG08], is used to show that the truncation in (18) does not effect the LSD too much. For any truncation level $\tau > 0$, let $\tau\mathcal{C}_n$ be a matrix with entries given by

$$\tau\mathcal{C}_n(j, k) := \mathcal{C}_n(j, k) 1_{\{|\mathcal{C}_n(j, k)| \leq \tau\}}. \quad (47)$$

Lemma 5.3 (Large deviation estimate for the rank of a truncation). *For every $\epsilon > 0$ and $\tau \gg 0$ (large enough depending on ϵ), there is a $\delta_{\epsilon, \tau} > 0$ such that*

$$\mathbf{P}(\text{rank}(\mathcal{C}_n - \tau\mathcal{C}_n)/n \geq \epsilon) \leq \exp(-\delta_{\epsilon, \tau} n).$$

Proof. Fix $\epsilon > 0$ and consider τ large enough (specified below). Define the events

$$U_{jn} := \{\text{there exists } k \text{ such that } k > j \text{ and } |\mathcal{C}_n(j, k)| > \tau\}$$

$$L_{jn} := \{\text{there exists } k \text{ such that } k < j \text{ and } |\mathcal{C}_n(j, k)| > \tau\}$$

and note that

$$\text{rank}(\mathcal{C}_n - \tau\mathcal{C}_n) \leq \sum_{j=1}^n (1_{U_{jn}} + 1_{L_{jn}}). \quad (48)$$

We split rows of the matrix along the diagonal to handle the dependence (due to the self-adjointness requirement) among the indicator random variables:

$$\begin{aligned}
\mathbf{P}(\text{rank}(\mathcal{C}_n - \tau \mathcal{C}_n) \geq 2n\epsilon) &\leq \mathbf{P}\left(\sum_{j=1}^n 1_{U_{jn}} \geq n\epsilon\right) + \mathbf{P}\left(\sum_{j=1}^n 1_{L_{jn}} \geq n\epsilon\right) \\
&\leq 2\mathbf{P}\left(\sum_{j=1}^n 1_{U_{jn}} \geq n\epsilon\right) \\
&\leq 2\mathbf{P}\left(\sum_{j=1}^n 1_{U_{1n}}^{(j)} \geq n\epsilon\right)
\end{aligned} \tag{49}$$

where $\{1_{U_{1n}}^{(j)}\}_{j=1}^n$ are independent copies of $1_{U_{1n}}$. The last step follows since the independent variables $\{1_{U_{jn}}\}_{j=1}^n$ are each stochastically dominated by $1_{U_{1n}}$.

Since the triangular array $\{\mathcal{C}_n(1, k), 1 \leq k \leq n\}_{n \in \mathbb{N}}$ satisfies (4),

$$\lim_{n \rightarrow \infty} \mathbf{P}(U_{1n}) = 1 - \exp\{-\Pi([\tau, \infty))\},$$

so we may choose τ large enough so that

$$\sup_n \mathbf{P}(U_{1n}) = p < \epsilon.$$

The lemma follows by applying a standard large deviation estimate for i.i.d. Bernoulli(p) random variables to the right side of (49). \square

This last lemma is used in conjunction with a metric which is compatible with weak convergence. Let

$$\|f\|_{\mathcal{L}} := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|} + \sup_x |f(x)|$$

Lemma 2.1 in [BAG08] says the following variant of the Dudley distance gives a topology which is compatible with weak convergence:

$$d_1(\mu, \nu) := \sup_{\|f\|_{\mathcal{L}} \leq 1, f \uparrow} \left| \int f d\mu - \int f d\nu \right|. \tag{50}$$

Moreover, Lidskii's estimate (see Eq. 8 in [BAG08]) implies

$$d_1(\mu_{\mathcal{C}_n}, \mu_{\tau \mathcal{C}_n}) \leq \frac{\text{rank}(\mathcal{C}_n - \tau \mathcal{C}_n)}{n}. \tag{51}$$

Proof of Theorems 1.1 and 1.3. Let us first state some simplifications for the task of showing that the LSD exists as a weak limit, almost surely.

First of all, by the Borel-Cantelli lemma and Lemma 5.1, it is enough to show weak convergence in expectation of $(\mathbf{E}\mu_{\mathcal{C}_n})_{n \in \mathbb{N}}$ to $\mathbf{E}\mu_{\mathcal{C}_\infty}$. Next, by exchangeability, it is enough to show weak convergence in expectation of the spectral measures associated to the basis vector e_1 . Finally, by Lemma A.2, it is equivalent to show convergence of the Cauchy-Stieltjes transforms of these expected spectral measures for each $z \in \mathbb{C}_+$ (the limit will be a probability measure since it is the spectral measure associated to a unit vector).

Choose a Lévy-Khintchine ensemble $(\mathcal{C}_n)_{n \in \mathbb{N}}$ and let $(\tau_m)_{m \in \mathbb{N}}$ be a sequence of positive truncation levels which go to infinity. For each truncation level τ_m , consider a new sequence of matrices

$(\tau_m \mathcal{C}_n)_{n \in \mathbb{N}}$ given by (47). Recalling our choice of h_n from Section 2, we also consider their modifications $(\tau_m \mathcal{C}_n^\sigma)_{n \in \mathbb{N}}$ (truncation occurs before modification).

Fix m . Each modified matrix sequence $(\tau_m \mathcal{C}_n^\sigma)_{n \in \mathbb{N}}$ satisfies the hypotheses of Proposition 4.5, thus the associated graphs have a PWIST($\sigma, \lambda_\Pi^{(m)}$) as their local weak limit as $n \rightarrow \infty$, where $\lambda_\Pi^{(m)}$ is the intensity λ_Π restricted to the set

$$(-\infty, -1/\tau_m] \cup [1/\tau_m, \infty).$$

The closure of the associated limiting operator is self-adjoint by Lemma 5.2. Moreover, by Proposition A.1 and the properties of the replacement procedure, we have for each $j \in \mathbb{N}$ that

$$\lim_{\epsilon \searrow 0} \lim_{n \rightarrow \infty} \sum_{k=1}^n \text{Var}(\mathcal{C}_n^\sigma(j, k) 1_{\{|\mathcal{C}_n^\sigma(j, k)| \leq \epsilon\}}) = 0 \quad (52)$$

which is equivalent to (38) since the entries $\mathcal{C}_n^\sigma(j, k) 1_{\{|\mathcal{C}_n^\sigma(j, k)| \leq \epsilon\}}$ have a real distribution which is symmetric for $\epsilon < \sigma$ (the truncation τ_m is unnecessary due to $1_{\{|\mathcal{C}_n^\sigma(j, k)| \leq \epsilon\}}$).

By the above considerations, we may use Theorem 4.2 and the argument below (41) to conclude Theorem 1.3 for each sequence $(\tau_m \mathcal{C}_n^\sigma)_{n \in \mathbb{N}}$. Thus, the expected LSD of $(\tau_m \mathcal{C}_n^\sigma)_{n \in \mathbb{N}}$, denoted by $\mathbf{E}\mu_{\tau_m \mathcal{C}_\infty^\sigma}$, is the expected spectral measure at e_\emptyset for the self-adjoint random conductance operator $\tau_m \mathcal{C}_\infty^\sigma$ associated to a PWIST($\sigma, \lambda_\Pi^{(m)}$).

Now take the local weak limit of the PWIST($\sigma, \lambda_\Pi^{(m)}$) graphs as $m \rightarrow \infty$. Since these graphs are truncations of a PWIST(σ, λ_Π), it is clear that their local weak limit is just a PWIST(σ, λ_Π). We may therefore apply Theorem 4.2 once more to conclude that the expected spectral measures at e_\emptyset of the PWIST($\sigma, \lambda_\Pi^{(m)}$) operators converge weakly to the expected spectral measure at e_\emptyset of a PWIST(σ, λ_Π) operator which we denote by $\mathbf{E}\mu_{\mathcal{C}_\infty^\sigma}$. Thus, for every $\epsilon > 0$ we can choose m large enough so that

$$d_1(\mathbf{E}\mu_{\tau_m \mathcal{C}_\infty^\sigma}, \mathbf{E}\mu_{\mathcal{C}_\infty^\sigma}) < \epsilon/3$$

and so that $\delta_{\epsilon, \tau_m} > 0$ in Lemma 5.3.

Eq. (22) and Propositions 3.2 and 3.3 show that the expected LSD for $(\tau_m \mathcal{C}_n)_{n \in \mathbb{N}}$ exists. Moreover, by Proposition 3.1, it is equal to $\mathbf{E}\mu_{\tau_m \mathcal{C}_\infty^\sigma}$. So we may choose n_0 large enough so that $n > n_0$ implies

$$d_1(\mathbf{E}\mu_{\tau_m \mathcal{C}_n}, \mathbf{E}\mu_{\tau_m \mathcal{C}_\infty^\sigma}) < \epsilon/3.$$

Lemma 5.3 and (51), show that we may finally choose n_1 large enough so that $n > n_1$ implies

$$d_1(\mathbf{E}\mu_{\mathcal{C}_n}, \mathbf{E}\mu_{\tau_m \mathcal{C}_n}) < \epsilon/3.$$

Combining the above, we have for all $n > \max(n_0, n_1)$,

$$d_1(\mathbf{E}\mu_{\mathcal{C}_n}, \mathbf{E}\mu_{\mathcal{C}_\infty^\sigma}) < \epsilon$$

and so the ESDs of $(\mathcal{C}_n)_{n \in \mathbb{N}}$ converge weakly in expectation (and thus a.s.) to $\mathbf{E}\mu_{\mathcal{C}_\infty^\sigma}$ which is the expected spectral measure at e_\emptyset of $\mathcal{C}_\infty^\sigma$ associated to a PWIST(σ, λ_Π).

The claim that $\mu_{\mathcal{C}_\infty}$ has bounded support if and only if Π is trivial, follows from the remark after the proof of Proposition \square

Proof of Corollary 1.2. The corollary follows from Theorem 2.1 in [BCC11b] since it is enough to show the existence of a limiting singular value distribution. We give a brief outline here and refer the

reader to [BCC11b] for more details. Let $\{\sigma_j\}_{j=1}^n$ denote the singular values of the n th matrix in the sequence $(\mathcal{A}_n)_{n \in \mathbb{N}}$ and define the symmetrized empirical measure

$$\sigma_{\mathcal{A}_n} := \frac{1}{2n} \sum_{j=1}^n (\delta_{\sigma_j} + \delta_{-\sigma_j}).$$

The idea is to associate a $2n \times 2n$ matrix \mathcal{B}_n to each \mathcal{A}_n by thinking of \mathcal{B}_n as an $n \times n$ matrix with entries given by the 2×2 matrices

$$\mathcal{B}_n(j, k) := \begin{bmatrix} 0 & \mathcal{A}_n(j, k) \\ \bar{\mathcal{A}}_n(j, k) & 0 \end{bmatrix}.$$

Through a permutation of entries, \mathcal{B}_n is similar to the block matrix

$$\begin{bmatrix} 0 & \mathcal{A}_n \\ \bar{\mathcal{A}}_n^* & 0 \end{bmatrix}$$

whose eigenvalues are $\pm \sigma_k(\mathcal{A}_n)$. Thus the ESD of \mathcal{B}_n is precisely equal to $\sigma_{\mathcal{A}_n}$, and we know that the LSD of $(\mathcal{B}_n)_{n \in \mathbb{N}}$ exists by Theorem 1.1. \square

Proof of Proposition 1.4. The proof is an application of the resolvent identity. For details, we refer the reader to Proposition 2.1 in [Kle98] or Theorem 4.1 in [BCC11a]. The latter proof works in our setting almost word for word. \square

A Some additional tools

Infinite divisibility. The following important set of criteria for convergence to an infinitely divisible law with characteristics (σ^2, b, Π) was found independently by Doeblin and Gnedenko (see Corollary 15.16 in [Kal02]). For $0 < h < 1$, define

$$\sigma_h^2 := \sigma^2 + \int_{|x| \leq h} x^2 \Pi(dx) \quad \text{and} \quad b_h := b - \int_{h < |x| \leq 1} x \Pi(dx).$$

Also, let $\bar{\mathbb{R}}$ be the one-point compactification of \mathbb{R} .

Proposition A.1 (Convergence criteria for triangular arrays). *Suppose $\{\mathcal{C}(n, k), 1 \leq k \leq n\}_{n \in \mathbb{N}}$ is a triangular array of random variables such that each row consists of i.i.d. random variables. The sum*

$$\sum_{j=1}^n \mathcal{C}(n, j)$$

converges in distribution to an $ID(\sigma^2, b, \Pi)$ random variable if and only if for any $0 < h < 1$ which is not an atom of Π ,

- $n\mathbf{P}(\mathcal{C}(n, 1) \in \cdot) \xrightarrow{w} \Pi$ on $\bar{\mathbb{R}} \setminus \{0\}$,
- $n\mathbf{E}(|\mathcal{C}(n, 1)|^2 1_{\{|\mathcal{C}(n, 1)| \leq h\}}) \rightarrow \sigma_h^2$,
- $n\mathbf{E}(\mathcal{C}(n, 1) 1_{\{|\mathcal{C}(n, 1)| \leq h\}}) \rightarrow b_h$.

From the Cauchy-Stieltjes transform to LSDs. The use of the Cauchy-Stieltjes transform in the context of random matrices dates back to Marčenko and Pastur [MP67]. Mainly, one obtains convergence of the ESDs of the random matrices $(C_n)_{n \in \mathbb{N}}$ by showing convergence of the Cauchy-Stieltjes transforms $(S_{\mu_{C_n}}(z))_{n \in \mathbb{N}}$ as defined in (40). The lemma given here is taken from Section 2.4 in [AGZ10].

The Cauchy-Stieltjes transform is invertible: For any open interval I such that neither endpoint is an atom of μ

$$\mu(I) = \lim_{y \rightarrow 0} \frac{1}{\pi} \int_I \Im S_\mu(x + iy) dx. \quad (53)$$

This uniquely determines the measure μ so that one then obtains the following result:

Lemma A.2 (Spectral convergence via Cauchy-Stieltjes transforms). *Suppose μ_n is a sequence of probability measures on \mathbb{R} and for each $z \in \mathbb{C}_+$, $S_{\mu_n}(z)$ converges to $S(z)$ which is the Cauchy-Stieltjes transform of some probability measure μ . Then μ_n converges weakly to μ .*

Proof. Let n_k be a subsequence for which μ_{n_k} converges vaguely to some sub-probability measure μ . For every $z \in \mathbb{C}_+$, $x \mapsto \frac{1}{x-z}$ is continuous and goes to 0 as $x \rightarrow \infty$. Thus one has $S_{\mu_{n_k}}(z) \rightarrow S_\mu(z)$ pointwise for each $z \in \mathbb{C}_+$. By the hypothesis, we have $S(z) = S_\mu(z)$. We then use (53) to see that every subsequence gives us the same limit which implies that μ_n converges vaguely to μ . But μ is a probability measure, thus we upgrade this to weak convergence. \square

References

- [AGZ10] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, volume 118. Cambridge University Press, 2010.
- [Ald92] David Aldous. Asymptotics in the random assignment problem. *Probability Theory and Related Fields*, 93(4):507–534, 1992.
- [Ald01] David Aldous. The $\zeta(2)$ limit in the random assignment problem. *Random Structures & Algorithms*, 18(4):381–418, 2001.
- [ALS07] Luigi Accardi, Romuald Lenczewski, and Rafal Salapata. Decompositions of the free product of graphs. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 10(03):303–334, 2007.
- [AS04] David Aldous and J. Michael Steele. The objective method: Probabilistic combinatorial optimization and local weak convergence. In *Probability on discrete structures*, pages 1–72. Springer, 2004.
- [BAG08] Gérard Ben Arous and Alice Guionnet. The spectrum of heavy tailed random matrices. *Communications in Mathematical Physics*, 278(3):715–751, 2008.
- [BC12] Charles Bordenave and Djalil Chafaï. Around the circular law. *Probability Surveys*, 9, 2012.
- [BCC11a] Charles Bordenave, Pietro Caputo, and Djalil Chafaï. Spectrum of large random reversible markov chains: heavy-tailed weights on the complete graph. *The Annals of Probability*, 39(4):1544–1590, 2011.
- [BCC11b] Charles Bordenave, Pietro Caputo, and Djalil Chafaï. Spectrum of non-Hermitian heavy tailed random matrices. *Communications in Mathematical Physics*, 307(2):513–560, 2011.
- [BG01] Michel Bauer and Olivier Golinelli. Random incidence matrices: moments of the spectral density. *Journal of Statistical Physics*, 103(1-2):301–337, 2001.

- [BG05] Florent Benaych-Georges. Classical and free infinitely divisible distributions and random matrices. *Annals of probability*, pages 1134–1170, 2005.
- [BGGM13] Florent Benaych-Georges, Alice Guionnet, and Camille Male. Central limit theorems for linear statistics of heavy tailed random matrices. *arXiv preprint arXiv:1301.0448*, 2013.
- [Bia97] Philippe Biane. On the free convolution with a semi-circular distribution. *Indiana University Mathematics Journal*, 46(3):705–718, 1997.
- [Bil86] P Billingsley. Probability and measure. *Wiley series in probability and mathematical statistics*, 1986.
- [BJN⁺07] Z. Burda, J. Jurkiewicz, M. Nowak, G. Papp, and I. Zahed. Random Lévy matrices revisited. *Phys. Rev. E*, 75, 2007.
- [BL10] Charles Bordenave and Marc Lelarge. Resolvent of large random graphs. *Random Structures & Algorithms*, 37(3):332–352, 2010.
- [BS01] Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. *Electronic Journal of Probability*, 6(10):1–13, 2001.
- [BS10] Zhidong Bai and Jack William Silverstein. *Spectral analysis of large dimensional random matrices*. Springer, 2010.
- [BV93] Hari Bercovici and Dan Voiculescu. Free convolution of measures with unbounded support. *Indiana University Mathematics Journal*, 42(3):733–774, 1993.
- [CB94] P. Cizeau and J.P. Bouchaud. Theory of Lévy matrices. *Physical Review E*, 50(3):1810, 1994.
- [DS07] R Brent Dozier and Jack W Silverstein. On the empirical distribution of eigenvalues of large dimensional information-plus-noise-type matrices. *Journal of Multivariate Analysis*, 98(4):678–694, 2007.
- [FZ97] Joshua Feinberg and A Zee. Non-hermitian random matrix theory: method of hermitian reduction. *Nuclear Physics B*, 504(3):579–608, 1997.
- [GL09] Adityanand Guntuboyina and Hannes Leeb. Concentration of the spectral measure of large wishart matrices with dependent entries. *Electron. Commun. Probab*, 14(334-342):4, 2009.
- [Gut98] Eugene Gutkin. Green’s functions of free products of operators, with applications to graph spectra and to random walks. *Nagoya Mathematical Journal*, 149:93–116, 1998.
- [HO07] Akihito Hora and Nobuaki Obata. *Quantum probability and spectral analysis of graphs*. Springer, 2007.
- [Kal02] Olav Kallenberg. *Foundations of modern probability*. Springer, 2002.
- [Kle98] Abel Klein. Extended states in the anderson model on the bethe lattice. *Advances in Mathematics*, 133(1):163–184, 1998.
- [KSV04] O. Khorunzhy, M. Shcherbina, and V. Vengerovsky. Eigenvalue distribution of large weighted random graphs. *Journal of Mathematical Physics*, 45:1648, 2004.
- [Küh08] Reimer Kühn. Spectra of sparse random matrices. *Journal of Physics A: Mathematical and Theoretical*, 41(29):295002, 2008.
- [Kyp06] Andreas E Kyprianou. *Introductory lectures on fluctuations of Lévy processes with applications*. Springer, 2006.
- [Mal11] Camille Male. The distributions of traffics and their free product. *arXiv preprint arXiv:1111.4662*, 2011.

- [Mal12] Camille Male. The limiting distributions of large heavy wigner and arbitrary random matrices. *arXiv preprint arXiv:1209.2366*, 2012.
- [MP67] Vladimir A. Marchenko and Leonid A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Matematicheskii Sbornik*, 114(4):507–536, 1967.
- [NS06] Alexandru Nica and Roland Speicher. *Lectures on the combinatorics of free probability*, volume 13. Cambridge University Press, 2006.
- [RB88] G.J. Rodgers and A.J. Bray. Density of states of a sparse random matrix. *Physical Review B*, 37(7):3557, 1988.
- [Res07] S.I. Resnick. *Heavy-tail phenomena: probabilistic and statistical modeling*, volume 10. Springer Verlag, 2007.
- [RS80] Michael C. Reed and Barry Simon. *Methods of Modern Mathematical Physics: Functional Analysis*, volume 1. Access Online via Elsevier, 1980.
- [Rya98] Øyvind Ryan. On the limit distributions of random matrices with independent or free entries. *Communications in Mathematical Physics*, 193(3):595–626, 1998.
- [Sos04] Alexander Soshnikov. Poisson statistics for the largest eigenvalues of Wigner random matrices with heavy tails. *Electron. Comm. Probab*, 9:82–91, 2004.
- [Ste02] J. Michael Steele. Minimal spanning trees for graphs with random edge lengths. In *Mathematics and Computer Science II*, pages 223–245. Springer, 2002.
- [Tao12] Terence Tao. *Topics in random matrix theory*, volume 132. AMS Bookstore, 2012.
- [Wig55] Eugene P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *The Annals of Mathematics*, 62(3):548–564, 1955.
- [WS80] Joachim Weidmann and Joseph Szücs. *Linear operators in Hilbert spaces*, volume 20. Springer New York, 1980.
- [Zak06] Inna Zakharevich. A generalization of Wigner’s law. *Communications in Mathematical Physics*, 268(2):403–414, 2006.